Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Ab-initio crystal structure analysis and refinement approaches of oligo *p*-benzamides based on electron diffraction data

Tatiana E. Gorelik,[a] Jacco van de Streek,[b] Andreas F. M. Kilbinger,[c] Gunther Brunklaus[d,e] and Ute Kolb[a]*

[a]Institute for Physical Chemistry, Johannes Gutenberg University Mainz, Jakob Welder Weg 11, 55128 Mainz, Germany, [b]Avant-garde Materials Simulation, Merzhauser Str. 177, 79100 Freiburg im Breisgau, Germany, [c]Chemistry Department, University of Fribourg, Chemin du Musée 9, CH-1700 Fribourg, Switzerland, [d]Max-Planck Institute for Polymer Research, Ackermanweg 10, 55128 Mainz, Germany, and [e]Institute for Physcial Chemistry, University of Muenster, Corrensstr. 28/30, 48149 Muenster, Germany

Correspondence e-mail: kolb@uni-mainz.de

Received 31 October 2011
Accepted 25 January 2012

Ab-initio crystal structure analysis of organic materials from electron diffraction data is presented. The data were collected using the automated electron diffraction tomography (ADT) technique. The structure solution and refinement route is first validated on the basis of the known crystal structure of tri-*p*-benzamide. The same procedure is then applied to solve the previously unknown crystal structure of tetra-*p*-benzamide. In the crystal structure of tetra-*p*-benzamide, an unusual hydrogen-bonding scheme is realised; the hydrogen-bonding scheme is, however, in perfect agreement with solid-state NMR data.

1. Introduction

Electron crystallography has long been used for crystal structure analysis: the first structure determination of an organic material from electron diffraction data dates back to 1936 (Rigamonti, 1936). Since then, structural reports based on electron diffraction data have appeared every now and then, but generally electron crystallography was not taken as a serious strategy for structure determination. Owing to the possibility of using a small probe size, electron diffraction is especially attractive for nano-crystalline materials and can be efficiently used in combination with X-ray powder diffraction since both techniques address the same type of sample.

The fact that electron diffraction has not been widely adopted as a structure analysis technique is mainly due to two major problems associated with electron diffraction data – *dynamical effects* in the intensities and a *low amount* of data. Dynamical effects appear because electrons interact with matter so strongly that even for very thin specimens secondary interaction between the diffracted beams occurs. The dynamical contribution can modify the intensities severely, making structure determination impossible. The situation is somewhat better for organic materials containing light scatterers, therefore, the major part of the electron diffraction structure analysis reports were done with organics (Kolb *et al.*, 2010). The interaction between the beams is especially enhanced when zonal data are collected: low-index crystallographic zones include a large number of simultaneously excited reflections allowing multiple paths for the beam interaction. All technical attempts to reduce the dynamical effects are based on reducing the multibeam interactions using *off-zone* data (precession electron diffraction; Own, 2005; oblique textured patterns; Vainshtein, 1964).

The low amount of data originates mainly from the way the data is traditionally collected – as a set of low index patterns. These patterns typically show high symmetry thus containing a large number of redundant reflections and thus a reduced number of independent reflections. Besides, a large number of

© 2012 International Union of Crystallography
Printed in Singapore – all rights reserved

Table 1
Lattice parameters of OPBA3 and OPBA4.

	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
OPBA3†	14.629	9.893	12.721	90	107.61	90
OPBA3 present study	14.70	9.88	12.68	89.624	107.08	90.742
OPBA4 primitive cell	25.50	17.27	5.32	90.17	95.47	76.80
OPBA4 centered cell	50.77	5.32	17.27	89.83	103.24	89.49
OPBA4 centered cell after DFT-D minimization	51.084	5.194	17.477	90	101.00	90

† According to Gorelik *et al.* (2010).

relatively high-index reflections between the positions of the collected zones is missing. Low completeness of the data comprises a severe problem for structure determination, therefore most of the structural work was done using a combination of electron diffraction data with other information – structural phases obtained from additional sources, energetic considerations or geometrical restraints (Kolb *et al.*, 2010).

A special complication concerning organic samples is the beam sensitivity of the material. Often the crystal decays during the time spent on the orientation of a zone pattern. Acquiring several zonal patterns from the same crystal through tilting is an even more complicated task. The use of these patterns for structure analysis is not straightforward either, since they may show different degrees of decay, additionally changing the intensity data.

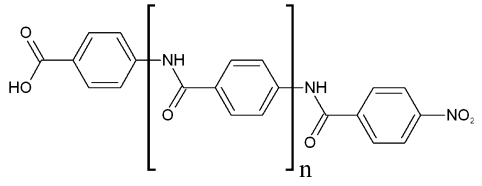
In order to overcome all the listed problems, the automated diffraction tomography (ADT) method was developed (Kolb *et al.*, 2007). An ADT experiment is performed in a transmission electron microscope (TEM), so electron diffraction patterns are recorded in transmission geometry. In ADT the electron diffraction data is collected sequentially while tilting a crystal around an arbitrary crystallographic axis in fine steps. Tilting around an arbitrary direction usually avoids zonal patterns and therefore reduces dynamical effects in the data. Fine sampling of the reciprocal space allows collection of a large number of reflections within the available tilt range. Tilting with steps of 1° gives good results and is practically reasonable. The use of a dedicated high-tilt, tomographic holder/stage setup permits tilting of up to ±60°, with a total tilt wedge of 120° out of the full 180° (which is 2/3 of the complete diffraction volume).

An automated ADT acquisition module was developed in collaboration with FEI for the TECNAI class of electron microscopes (Kolb *et al.*, 2007). Electron diffraction data is collected in nano-diffraction mode using a small (10 µm) condenser aperture. By focusing the objective lens, an illuminated spot with a size of 50–100 nm can be produced at the sample with a quasi-parallel beam. To ensure that the diffraction data is collected from the same part of the sample, crystal tracking routines are used. The routines are adopted from the direct space tomography approach and include a holder/stage setup calibration eliminating rough movements of the stage and fine tracking routines based on cross-correlation of the neighboring images. In the ADT acquisition module crystal imaging is done in the scanning transmission electron microscopy (STEM) mode, therefore, STEM images are cross-correlated. Once a suitable crystal has been selected, the complete automated sequence can be expressed as tilt → crystal tracking → diffraction pattern acquisition. The data is stored as a stack of diffraction patterns with relevant header information in MRC format. A dedicated software package is then used to process the tilt series and extract intensity data sets (*ADT3D*, Nanomegas, Belgium).

Although ADT was initially developed for electron-beam-sensitive materials, it also turned out to be very efficient for inorganic samples, and recently a number of inorganic crystal structures solved *ab initio* from ADT data were reported (Kolb *et al.*, 2011).

The structure analysis of two members of a homologous series of oligo-*p*-benzamides is described in this paper: OPBA3 and OPBA4 (Fig. 1). Oligo-*p*-benzamides are important building blocks for supramolecular chemistry (Abbel, Frey *et al.*, 2005; Abbel, Schleuß *et al.*, 2005), self-organized due to extended hydrogen-bond systems and π-stacking interaction. Rigid fiber-like structures were reported for copolymers of hepta-*p*-benzamide with polyethylene glycol (PEG; König *et al.*, 2007). It was proposed that the structures consist of rigid OPBA cores held together by hydrogen bonds and π-stacking interactions surrounded by the PEG coils. The exact arrangement of the core, however, was not known. As an initial step towards the molecular packing of the higher oligomers, the crystalline structures of the shorter members were studied.

However, even the shortest members of the series turned out to be poorly crystalline. Their structures, solved by a combination of zonal electron diffraction data and X-ray powder diffraction, were reported recently (Gorelik *et al.*, 2010). All recrystallization attempts for the next homolog – tetra-*p*-benzamide – did not improve the crystallinity. The X-ray powder profile contained only a few broad peaks. In this manuscript the known structure of OPBA3 was used to validate the structure analysis and refinement routine based on ADT data. The structure of OPBA4 was not known hitherto and was solved *ab initio* from ADT data.



OPBA3: n=1
OPBA4: n=2

Figure 1
OPBA3 and OPBA4 molecules.

2. Experimental

OPBA3 and OPBA4 were synthesized using the procedure reported in Abbel, Frey *et al.* (2005).

Samples for TEM investigations were prepared by suspending the material in *n*-hexane in an ultrasonic bath. A drop of the suspension was placed on a carbon-film-coated copper grid and dried in air. TEM investigations were performed with a TECNAI F30 transmission electron microscope equipped with a field emission gun operating at 300 kV. Electron diffraction data were collected using a high-tilt tomography holder (FISCHIONE) within a $\pm 60^\circ$ tilt range with a tilt step of 1°. The diffraction data were recorded on a 1k GATAN CCD camera.

Automated diffraction tomography (ADT) data were collected using a dedicated acquisition module (Kolb *et al.*, 2007). As the crystal was tilted the beam was moved across the crystal to ensure the diffraction patterns were collected from a fresh, possibly not previously exposed area. The data was collected using an electron dose rate of 0.2 e $Å^{-2}$ s^{-1} and the exposure time for each frame was 5 s. The total electron dose received during data collection was 0.2 e $Å^{-2}$ s^{-1} × 5 s per frame × 121 frames = 121 e $Å^{-2}$. This dose was a distribution over a certain area of a crystal, thus the mean value is even lower.

ADT data were processed using the *ADT3D* package (Nanomegas, Belgium). *Ab-initio* structure solution was performed in *SIR*2008 (Burla *et al.*, 2005). Structure refinement based on electron diffraction data was carried out using *SHELX* (Sheldrick, 2008). Geometry optimization using energy minimization was carried out using the *COMPASS* force field (Materials Studio, Accelrys Inc., USA) using charges assigned by the force field.

2.1. X-ray powder diffraction

Data was collected in reflection geometry on a Siemens D500 diffractometer using U = 40 kV, I = 35 mA and a Ge(111) monochromator (Cu $K\alpha_1$, λ = 1.54056 Å). The data was recorded from 2 to 35° with a step of 0.02° and the integration time at each step was 60 s.

2.2. Solid-State NMR

Proton solid-state NMR data were recorded on a Bruker Avance 700 spectrometer at a spinning frequency of $\sim$ 30 kHz, typical $\pi/2$ pulse lengths of 2.5 µs and recycle delays of 5 s. The back-to-back (BaBa) recoupling sequence was used to excite and reconvert double-quantum coherences (Saalwächter *et al.*, 2001). For each of the 64 t_1 slices, 64 transients were coadded, applying the States TPPI method for phase-sensitive detection (Marion *et al.*, 1989). All ^{13}C cross-polarization and magic angle spinning (^{13}C CP-MAS) spectra were collected at 125.77 MHz (Bruker Avance 500 spectrometer), with a CP contact time of 3 ms coadding 2048 transients. The experiments were carried out using a standard Bruker 2.5 mm double-resonance MAS probe spinning at 25 kHz, with a typical $\pi/2$-pulse length of 2.5 µs and a recycle delay of 5 s. In addition, ^{15}N CP-MAS spectra were collected at 30.4 MHz (Bruker Avance-II 300 spectrometer), with a CP contact time of 9 ms coadding 8192 transients. The experiments were carried out using a standard Bruker 4 mm double-resonance MAS probe spinning at 5 kHz, with a typical $\pi/2$ pulse length of 4 µs and a recycle delay of 15 s. All ^1H and ^{13}C spectra are referenced with respect to tetramethyl silane (TMS) using adamantane as a secondary standard (1.86 p.p.m. for ^1H and 29.46 p.p.m. for ^{13}C). All spectra were collected at room temperature. ^{15}N spectra were referenced with respect to solid ^{15}N-NH_4NO_3 (−358.4 p.p.m.; Goward *et al.*, 2001).

^1H–^1H double-quantum (DQ) MAS NMR spectrum of OPBA4 at 700.1 MHz and 29762 Hz MAS, acquired under the following experimental conditions: $\tau_{(exc.)}$ = 33.6 µs, 64 t_1 increments at steps of 33.6 µs, relaxation delay 5 s, 64 transients per increment; 16 positive contour levels between 2% and 83% of the maximum peak intensity were plotted. The F_2 projection is shown at the top.

2.3. Dispersion-corrected density functional theory (DFT-D) calculations

The crystal structures of OPBA3 and OPBA4, determined from electron diffraction data, were optimized using the program *GRACE* (Neumann & Perrin, 2005; Neumann, 2011), which uses *VASP*4.6 (Kresse & Furthmüller, 1996; Kresse & Hafner, 1993; Kresse & Joubert, 1999) for single-point pure DFT calculations. The generalized gradient approximation (GGA) with the Perdew–Wang 91 (Wang & Perdew, 1991) exchange-correlation functional was used, with standard projector-augmented wave (PAW) potentials. The plane-wave cut-off energy was 520 eV and the *k*-point spacing was approximately 0.7 $Å^{-1}$. The settings for the DFT calculations as well as a full description of the dispersion correction are given in Neumann & Perrin (2005). The r.m.s. Cartesian displacement of the non-H atoms was calculated as described in van de Streek & Neumann (2010). Values up to 0.25 Å indicate that the structure is correct, whereas values greater than 0.30 Å point to an incorrect structure. The structures were energy minimized in three steps, gradually releasing more degrees of freedom to improve the odds of the minimization converging to the nearest minimum. In the final step

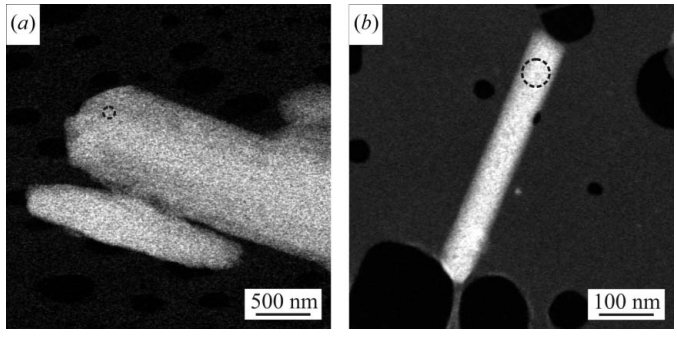


Figure 2
Crystals of (*a*) OPBA3 and (*b*) OPBA4. The size of the beam used for data collection is marked by a black dashed circle.

Table 2
Intensity data sets and structure solution parameters of OPBA3 and OPBA4.

	OPBA3	OPBA4
Crystallographic information		
Space group	$P2_1/c$ (Z = 4)	$C2/c$ (Z = 8)
Cell volume (Å^3)	1755	4552
Calculated density (g cm^{-3})	1.534	1.531
Intensity data sets		
Collected reflections	9185	12 805
Independent reflections	3078	3576
Resolution range (Å)	8.07–0.78	8.58–0.80
$R_{sym}(F)$	0.346	0.306
Completeness (%)	81	77
B_{iso}/U_{iso}	2.55/0.03	2.98/0.04
Structure solution (*SIR*)		
Final residual value $R(F)$	0.43	0.46
Atoms found (non-H)	30 (30)	38 (39)
Structure refinement (*SHELX*)		
Final residual value (restraints) $R1(F)$	0.58	0.65

all the degrees of freedom, including the unit-cell parameters, are relaxed and only the experimental space group is imposed.

3. Results and discussion

3.1. OPBA3

An electron diffraction tilt series was collected, automatically with a 1° tilt step in the tilt range of ±60°, from the crystal shown in Fig. 2(*a*). The position of the nano-beam was slightly changed during tilting so that the total electron dose was distributed over a large area of the crystal.

The reconstruction of the diffraction volume of the OPBA3 tilt series is shown in Figs. 3(*a*) and (*b*). When projected along the (100) direction, rows of extinctions along the **c*** axis are seen. These correspond to extinctions due to a *c*-type glide plane. This projection does not show serial extinctions along the monoclinic **b*** axis. The unit-cell vectors were found using automated routines (Kolb *et al.*, 2008; see Table 1).

3.1.1. Structure solution. An intensity data set was extracted using published methods (Mugnaioli *et al.*, 2009). Table 2 summarizes the characteristics of the data set. The structure was solved using the direct methods approach as implemented in *SIR*2008 (Burla *et al.*, 2005) in the space group $P2_1/c$. The first 30 potential peaks could be assigned to all non-H atoms of the OPBA3 molecule. The molecular geometry of the structure was found correctly, the phenyl rings were somewhat distorted (Fig. 4*a*). The r.m.s. Cartesian displacement of the atoms compared with the previously reported structure is 0.17 Å with a maximum displacement of 0.32 Å (C3 atom in a phenyl ring).

3.1.2. Refinement based on electron diffraction data: *SHELX*. Structure refinement was carried out in *SHELX* (Sheldrick, 2008). Refinement with no restraints applied was stable but did not change the geometry of the model significantly. Therefore, the final refinement procedure was carried out using constraints on the geometry of the phenyl rings (afix 66), planarity restraints (flat) on the nitro group, on acetamido groups and on the carboxy group. Additionally, for the nitrogen–

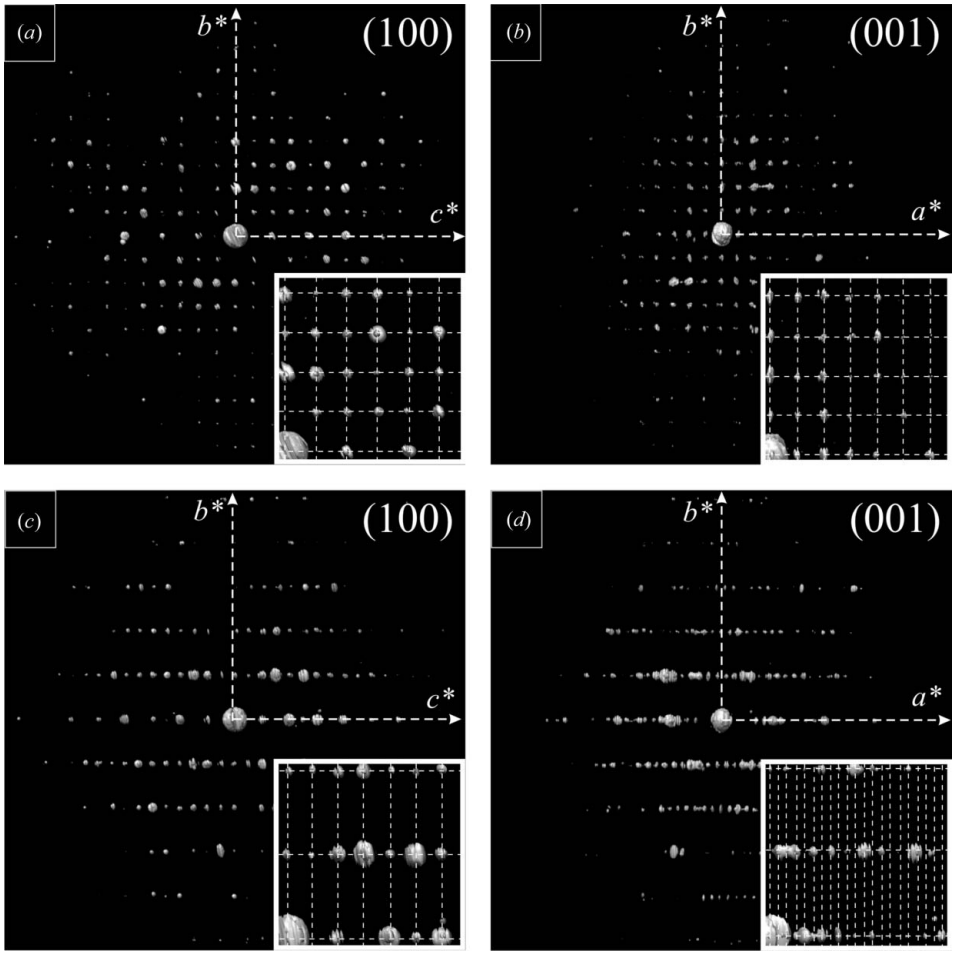


Figure 3
Views of the reciprocal volume along main directions of (*a*), (*b*) OPBA3 and (*c*), (*d*) OPBA4; reflection extinctions due to a glide plane are visible along the *c** axis in (100) projections (*a*) and (*c*), integral extinctions due to C-centering of OPBA4 are evident in (*d*). The two directions *a** and *c** are not in the plane of the figure. Inserts in the bottom-right corners: enlarged central parts of the projections overlaid with the reciprocal lattice.

Table 3
Match between the previously published OPBA3 structure and structural models obtained in different ways.

	SIR	*SIR/SHELX*	*SIR/COMPASS*	*SIR/SHELX/COMPASS*
Molecules matched (out of 15)	15	9	15	15
R.m.s. Cartesian displacement (Å)	0.17	0.29	0.19	0.18

oxygen bond and the carbon–oxygen bond in the nitro group the same distance restraints (sadi) within two acetamido groups were used.

The refinement converged to an $R1(F)$ of 0.58. The displacement parameters were refined isotropically; after the refinement three C atoms had negative displacement parameters. Interestingly, the residual peaks in the Fourier difference map showed two additional phenyl rings, which could be interpreted as an alternative position of the molecule pointing to positional disorder within the structure.

Although the final residual factors are high, the refined structure is very similar to that previously reported, showing an r.m.s. atomic displacement of 0.29 Å. The maximum displacement of 0.54 Å was observed for an O atom of the nitro group, although the thermal factor of the atom was reasonable.

3.1.3. Refinement using energy considerations. For OPBA3, although the phenyl rings did not have a perfectly planar geometry, the structure solution reproduced the expected structure quite well. In order to adjust the local bond geometry, the geometry of the complete structure was adjusted using energy minimization with the *COMPASS* force field. Thus, the energy minimization step was used instead of the refinement procedure.

In principle, there is no guarantee that the reliability of force-field calculations is transferable: the *COMPASS* force field may prove to work well for one compound, but may deliver a false geometry for compounds containing other structural features. Therefore, prior to the use of energy minimization as a substitute for the structure refinement procedure, the force field has to be validated for the given type of molecules. In the case of OPBA, where the members of the homologous series are chemically very similar, showing that the force field performs well on shorter oligomers OPBA2 and OPBA3 should make it also applicable to higher members.

All tested OPBA molecules were treated as flexible moieties. The lattice parameters were not optimized at this stage. The application of the *COMPASS* force field to the OPBA2 crystal structure was tested by Gorelik *et al.* (2010). In fact, the crystal structure of OPBA2 was solved using the *COMPASS* energy term. The final energy of the optimized structure of OPBA3 was -334 kJ mol^{-1}. The molecular geometry did not change significantly compared with the structure solved from X-ray powder diffraction data (Gorelik *et al.*, 2010). The average atomic displacement between the assolved and optimized structures was 0.17 Å. Thus, the geometry optimization based on *COMPASS* energy minimization can in principle be used as a substitute for the refinement procedure for the OPBA series.

After the geometry optimization procedure was applied to the *SIR* structure solution of OPBA3 (*SIR/COMPASS*), the total energy was -334 kJ mol^{-1} for the unit cell with $Z = 4$, giving -334.9 kJ mol^{-1} per molecule. The fit to the previously reported structure (Gorelik *et al.*, 2010) was excellent: the average atomic displacement was 0.16 Å with a maximal displacement of 0.35 Å for an O atom of the carboxylic acid group.

COMPASS geometry optimization on the OPBA3 crystal structure after refinement with *SHELX* (*SIR/SHELX/COMPASS*) was also stable upon energy minimization and delivered the same structure as the previous model (*SIR/COMPASS*). The average atomic displacement compared with the previously reported structure was 0.19 Å with a maximal displacement of 0.35 Å for an O atom of the carboxylic group. The resulting molecular conformation and overlay with the known crystal structure are shown in Figs. 4(*b*) and (*c*).

3.1.4. Structure validation. The classical structure solution route includes structural model refinement against the experimental diffraction data (*SHELX* refinement). The final residual factors for the *SIR* structure solution and *SHELX* refinement based on electron diffraction data are very high, and although the complete molecule was determined correctly these values would not be considered to be acceptable in X-ray crystallography. The geometry optimization using *COMPASS* resulted in a structure very similar to the expected

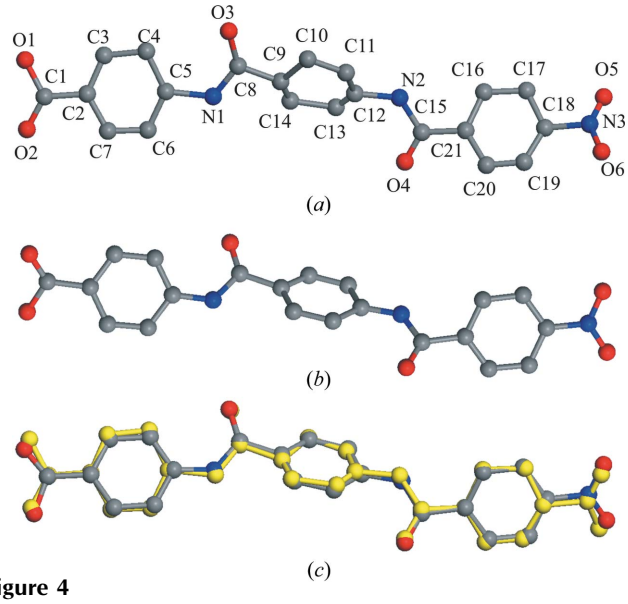


Figure 4
Molecular geometry in the crystal structure of OPBA3: (*a*) *SIR* structure solution from ADT data; (*b*) the molecule conformation after *SHELX* refinement followed by *COMPASS* energy minimization; (*c*) overlay of *SIR/SHELX/COMPASS* conformation and earlier determined structure as reported in Gorelik *et al.* (2010)

Table 4
Internal consistency check of OPBA3 and OPBA4 structures *versus* the structure model obtained after DFT-D minimization (see the text).

	Known structure XRPD (Gorelik *et al.*, 2010)	*SIR*	*SIR/SHELX*	*SIR/COMPASS*	*SIR/SHELX/COMPASS*
OPBA3					
Molecules matched (out of 15)	15	15	15	15	15
R.m.s. Cartesian displacement (Å)	0.18	0.22	0.27	0.23	0.22
OPBA4					
Molecules matched (out of 15)	–	3	–	15	15
R.m.s. Cartesian displacement (Å)	–	0.37	–	0.21	0.21

model, so principally this kind of structure optimization can be used as an alternative to the classical refinement.

The match between the known structure and all refined models was evaluated using the packing similarity algorithm implemented in *Mercury* (Macrae *et al.*, 2008; Chisholm & Motherwell, 2005). The algorithm returns the number of molecules that matched out of the default number (15) tried, and the non-H r.m.s. Cartesian displacement for the molecules that matched; Table 3 lists the results. For the *SIR* solution, the *SIR*/*SHELX*/*COMPASS* and *SIR*/*COMPASS* structures, all 15 molecules matched and the r.m.s. displacements are very low, indicating that these structures are '*good*' models. For the *SIR*/*SHELX* structure, only nine molecules out of 15 matched, and these nine already had a relatively high displacement of 0.29 Å, classifying this model as '*slightly worse than the others but still reasonable*'.

The validation described above is applicable exclusively to crystal structures that are already known. For unknown structures a method evaluating internal structure consistency without using any experimental information is necessary. Recently, the application of dispersion-corrected density functional theory (DFT-D) to validate experimental molecular structures was reported (van de Streek & Neumann, 2010) as described above. The r.m.s. Cartesian displacement between the experimental model and the result from the final minimization step (all degrees of freedom, including unit-cell parameters, are free) can be used as a criterion for structure correctness. From a validation against a test set of 241 single-crystal structures, it was found that an r.m.s. Cartesian displacement greater than 0.30 Å means that the structure is probably wrong (van de Streek & Neumann, 2010). A value smaller than 0.25 Å generally means that the structure is correct.

According to this validation strategy, the OPBA3 structure solved earlier from X-ray powder diffraction data is correct (all molecules match with an r.m.s. Cartesian displacement of 0.18 Å, see Table 4). Interestingly, the *SIR* structure solution, although having strong distortions in bond geometry, is still clearly correct (r.m.s. Cartesian displacement of 0.22 Å); the *SHELX* refined and *COMPASS* minimized structures do not show significant differences.

3.1.5. Refinement strategies. For OPBA3 the high value of the residual factor after the refinement procedure approaches that expected for a random structure (83% for a centrosymmetric structure built of one type of atom; Wilson, 1950). Experimental electron diffraction intensities are affected by a dynamical contribution, although this is thought to be low for *off-zone* data acquisition and light scatterers. Nevertheless, the dynamical component of ADT data has not been systematically studied yet. The effect of crystal mosaicity plausible for organic material should also modify the diffraction patterns. Finally, practical issues such as the different degree of decay imprint in diffraction patterns for beam-sensitive crystals and the wrong reflection indexing due to crystal bending strongly distort the experimental data. An interplay of these factors accounts for the high reliability index after the refinement procedure. The stability of the refinement procedure can be used as an additional sign of the consistency between the model and the experimental data, but cannot be expected to improve the model, except for correcting the local geometry when corresponding constraints are applied.

Another problem which can hamper the refinement procedure is the lack of data. For a monoclinic Laue class during ADT data acquisition within the $\pm 60^\circ$ wedge, typically around 80% of the reflection completeness can be achieved.

Refinement of the structure models using diffraction data as in single-crystal X-ray analysis is currently not very efficient for electrons. The refinement procedure is optimized for X-rays and cannot be directly transferred to electron diffraction data having its own peculiarities such as dynamical effects and reduced completeness. Nevertheless, the refinement can give a general idea of the structure model consistency and stability of the solution.

Force field structural model geometry optimization (*COMPASS*) turned out to be a more reliable approach. *COMPASS* optimization of the *SIR* structure solution and *SIR*/*SHELX* model converged practically to the same structure. Therefore, in order to obtain a final structure model it is sufficient to perform the force field geometry optimization directly on the *SIR* solution excluding the *SHELX* step.

3.2. OPBA4

The crystals of OPBA4 are much smaller than those of OPBA3 and all have a pronounced needle-like morphology. An electron diffraction tilt series was collected, using a beam size as indicated in Fig. 2(*b*). OPBA4 crystals appeared to be more sensitive to electron irradiation than OPBA3.

Figs. 3(*c*) and (*d*) show projections of the reconstructed reciprocal space of OPBA4. A row of zonal extinctions is seen along a main direction, interpreted as a *c*-glide plane in the crystal structure. When viewed in the **c*** direction, the

reflections are arranged in a checkerboard pattern, according to an integral extinctions law corresponding to a C-centered lattice. First the primitive lattice vectors were found, which were then recalculated into the centered unit cell (see Table 1).

ADT unit-cell vectors carry stochastic (precision) and systematic (accuracy) errors. The lattice vectors are calculated in difference vector space using clustering routines (Kolb *et al.*, 2008). Therefore, the precision of the determination of the vectors is effectively the relative size of the clusters. The cluster size is influenced most by the excitation error (Hirsch *et al.*, 1977) and effects associated with the extended shape of reflections (thin foil effect and disorder). For electron-beam-sensitive materials, when the electron dose has to be distributed over a large area, crystal bending can enlarge the cluster size significantly. Nevertheless, ADT unit-cell parameters are usually internally highly consistent, which can be judged for instance by the monoclinic α and γ angles which are usually very close to 90° when experimentally determined from ADT data.

The accuracy of the ADT lattice parameter determination is defined by the electron diffraction camera length calibration. Additional focusing due to the use of the nano-diffraction mode changes the effective camera length and therefore introduces a scaling factor for the lattice parameters. This factor can be calibrated against the diffraction lens currently used for the additional focusing. The final accuracy is usually around 2% (Kolb *et al.*, 2011).

Correct lattice parameters are essential for the effective use of minimization procedures. For OPBA4 the quality of the X-ray powder diffraction (Fig. 5) pattern did not allow refinement of the unit-cell parameters. Therefore, the density of OPBA3 (Z = 4) was used to scale the length parameters of OPBA4 (Z = 8). The resulting lattice parameters are monoclinic: a = 50.77, b = 5.32, c = 17.27 Å, β = 103.24°. These parameters were used for structure solution. Subsequent energy minimization of the obtained model led to a slight modification of these values, so the final lattice parameters used were a = 51.084, b = 5.194, c = 17.477 Å, β = 101.00° (see Table 1).

The space-group symmetry of the crystal structure includes C-centering and a c-glide plane. Two space groups satisfy these extinction rules: $C2/c$ and the non-centrosymmetric Cc. For the sake of simplicity, the initial structure solution attempts were undertaken in the space group $C2/c$.

3.2.1. Structure solution. The intensities of the reflections were extracted using the primitive unit-cell vector set, and then the indices were transformed to the centered lattice. A summary of the data set characteristics and structure solution is presented in Table 2. Atoms found by *SIR* were connected manually into an OPBA4 molecule. The only missing atom was one O atom of the NO_2 group. Several ghost peaks were present in the structure; the model was cleaned from ghosts and atomic species were assigned correctly. The resulting structure is presented in Fig. 6(*a*). All four phenyl rings, although they appear strongly distorted, are completely resolved. The missing O atom was added based on the idealized NO_2 group geometry. Interestingly, already at this stage it is seen that the COOH groups of adjacent molecules do not approach to form a synthon. Instead, the COOH group appears close to the NO_2 group.

The middle part of the OPBA4 molecule appears to be nicely resolved and to have a more correct geometry than the edges of the molecule. For the structure this means that there are bands in the structure, with a periodicity of 1/2a, which are not clearly resolved. It turned out that the low-resolution reflections such as (200) carrying structural information of these regions are not included in the data set. Due to the strong preferred orientation of the crystals on the TEM sample grid, certain directions cannot be reached. The reflections associated with long crystallographic axes are especially difficult to sample. In some cases crystal sectioning may help to record the missing orientation, and ultra-

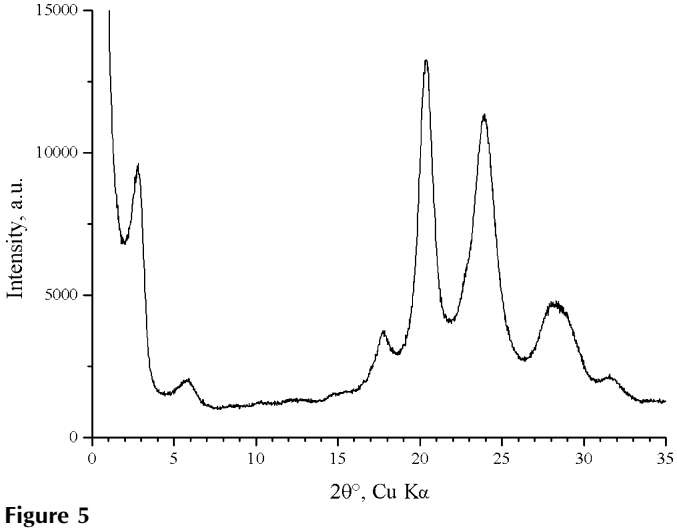


Figure 5
Powder X-ray profile of OPBA4.

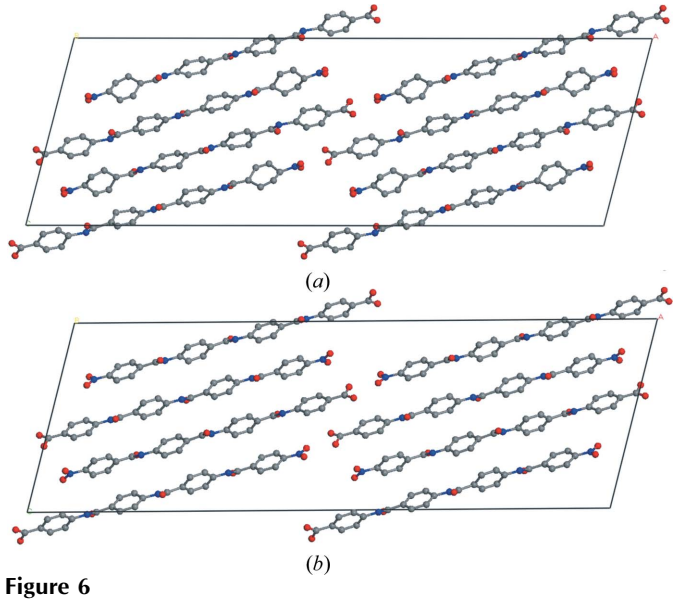


Figure 6
Crystal structure of OPBA4: (*a*) *SIR* structure solution from ADT data; (*b*) *SIR*/*COMPASS* structure.

microtomy was already successfully used to reach the desired orientation. However, working with cuts of beam-sensitive materials is very difficult in practice because the cut crystals are particularly small and damaged, so there is not enough homogeneous crystal area to use the dose distribution strategy during the data collection. Therefore, the missing low-resolution reflection is a special concern for *ab initio* structure solution of beam-sensitive materials and new strategies to resolve this problem need to be developed.

The structure solution was also tried using the non-centrosymmetric space group *Cc*. In this case the terminal phenyl rings of the molecule were not resolved. As the middle part of the molecule (three phenyl rings) had the same conformation as for the structure solution in *C*2/*c*, a centrosymmetric space group was presumed.

3.2.2. Structure refinement and validation. DFT-D structure optimization was performed on the *SIR* structure solution model. Although generally showing a very similar packing, the minimized structure showed a relatively poor match with the *SIR* structure solution – only three molecules were matched (using the default settings) with the r.m.s. Cartesian displacement of 0.37 Å.

The refinement in *SHELX* was performed using the same type of constraints/restraints as those applied for OPBA3 refinement. The terminal parts of the molecule significantly declined from their original positions. The displacement factors of eight atoms became negative. The DFT-D optimized structure showed even worse agreement than that for the *SIR* solution. After the *SHELX* refinement the two terminal phenyl rings were significantly shifted (more than 0.5 Å). The refinement was carried out using severe restraints on the geometry. The diffraction data poorly represented terminal parts of the molecule (see the structure solution description), therefore, the refinement of these parts was mostly governed by the restraints. Obviously, for a flexible molecule like OPBA4 this situation resolved into a distortion of the structure. Moreover, there were already hints of positional disorder as seen for OPBA3. Larger and more flexible molecules having stronger tendency towards non-regular packing can be expected. Thus, possible disorder in the structure can additionally hamper the refinement procedure.

The COMPASS force-field-based geometry minimization procedure was first applied directly to the *SIR* solution and to the *SIR*/*SHELX* structure. As expected, there was no significant difference in the result: both converged to the same structure (with an r.m.s. displacement of 0.003 Å between them) having −481 kJ mol^{-1}. The resulting structure is shown in Fig. 6(*b*). *COMPASS* minimization corrected the distortion and resulted in a structural model which matched the DFT-D minimized structure quite well (r.m.s. Cartesian displacement of 0.21 Å for all 15 molecules, see Table 4).

A rather unusual feature of the structure is the formation of a hydrogen bond between the carboxylic acid and the nitro group (Fig. 7) instead of the typical COOH dimer synthon (Gilli & Gilli, 2009). This probably happens owing to sterical hindrance when packing relatively rigid long-chain molecules into a crystal.

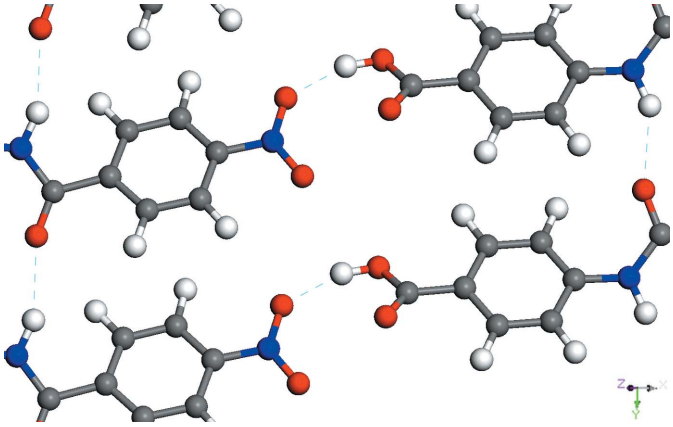


Figure 7
Hydrogen bond between the nitro and carboxylic group in the crystal structure of OPBA4.

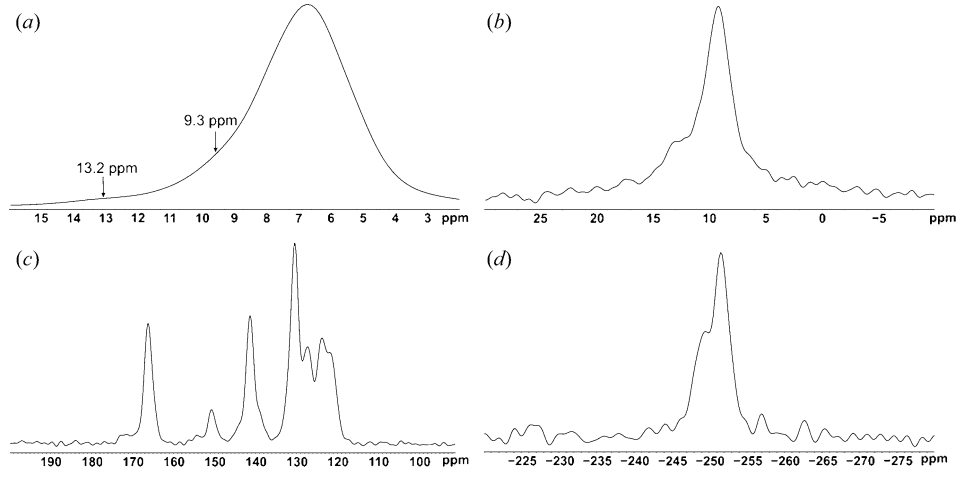


Figure 8
NMR data of OPBA4: (*a*) ^{1}H MAS NMR spectrum; (*b*) ^{2}H MAS NMR spectrum; (*c*) ^{13}C CP-MAS NMR; (*d*) ^{15}N CP-MAS NMR.

In the crystal structure the OPBA4 molecules are not planar. The four phenyl rings are twisted with respect to each other so that the dihedral angle between the least-squares planes through the non-H atoms of the first (starting from the carboxylic acid side) and the second ring is approximately 59.5°, the angle between the second and the third ring is −56.9°, and between the third and the fourth phenyl ring is 56.8°. In other words, the phenyl ring orientations are alternating along the molecule: the first ring is almost parallel to the third and the second almost parallel to the fourth, with a 60° rotation between the two sets.

3.2.3. Solid-state NMR studies. Based on an intensity distribution of $\sim$1:3:16, the ^{1}H MAS NMR spectrum of OPBA4 (Fig. 8*a*) reflects the molecular structure of OPBA4 exhibiting a weak peak at 13.2 p.p.m., a 'shoulder' at 9.3 p.p.m. and a rather broad, asymmetric signal centered at $\sim$ 6.8 p.p.m. The latter signal can be assigned to the aromatic protons of OPBA4, while the peak at 13.2 p.p.m. is attributed to the COOH proton. Notably, the NH proton can be identified at $\sim$ 9.3 p.p.m. in the ^{1}H MAS NMR spectrum, but is similar to the COOH proton more clearly revealed in the corresponding ^{2}H MAS NMR spectrum (Fig. 8*b*). Note that in ideal cases where the hydrogen-bonding pattern remains unchanged, the resonances in both ^{1}H and ^{2}H MAS NMR spectra are comparable within $\pm$0.1 p.p.m. (Brunklaus *et al.*, 2009).

The corresponding ^{13}C CP-MAS (Fig. 8*c*) spectrum of OPBA4 exhibits at least seven different ^{13}C resonances, which implies that the asymmetric unit comprises one molecule with no internal symmetry. In the carbonyl unit region the expected CONH signal is nicely resolved at 166.4 p.p.m., while the rather weak COOH signal shows some splitting at 171. 6 and 173.1 p.p.m., respectively, which possibly indicates local disorder. Similarly, the signal at 141.4 p.p.m. (assigned to the *ipso*-carbon of the phenyl ring attached to NH) may reflect broadening due to residual dipolar couplings (RDCs) between ^{13}C and ^{14}N spin pairs that comprise the C—N bond. The magnitude of such splitting, however, scales inversely with the static magnetic field strength of the NMR experiment and is thus fairly small at intermediate field strengths such as 11.7 T (Gan, 2006). In principle, line broadening due to ^{13}C–^{14}N RDCs can be either decoupled artificially *via* selective pulse irradiation employing a non-standard ^{14}N, ^{13}C and ^{1}H triple-resonance MAS NMR probe (Orr & Duer, 2006), or self-decoupled due to sufficiently fast quadrupolar relaxation (Olivieri, 1989). A detailed analysis of the ^{14}N quadrupolar coupling tensor (*e.g.* of amide bond N atoms) and the resulting RDCs, however, may be performed based on either (field-dependent) ^{13}C-detected wideline ^{14}N MAS NMR (Gan, 2008) or two-dimensional heteronuclear multi-quantum ^{13}C–^{14}N MAS NMR (Antonijevic & Halpern-Manners, 2008), which is beyond the scope of this work. Notably, further line broadening due to local structural disorder cannot be excluded. Since the closest proton contact of the amide carbonyl unit comprises aromatic protons, a slight change in the tilt angle of the phenyl ring with respect to the amide plane may contribute to the observed signal splitting.

Also, OPBA4 displays two peaks in the ^{15}N CP-MAS spectrum (Fig. 8*d*) at $-$248.4 and $-$254.8 p.p.m., respectively, where the corresponding integrated area ratio of 1:2 corroborates an assignment to the amide-NH units; the line width of the peaks is fairly narrow (about 92 Hz). Notably, unlike in the previous case of OPBA2 (Gorelik *et al.*, 2010), no peak attributable to the nitro group has been detected, suggesting local dynamics on the ms scale perturbing the ^{1}H–^{15}N polarization transfer (Krushelnitsky *et al.*, 2002).

A correlation between protons participating in different hydrogen-bond motives can be mapped by a two-dimensional ^{1}H–^{1}H double-quantum correlation spectrum (Khan *et al.*, 2010; Bolz *et al.*, 2008). In such a two-dimensional experiment, double-quantum coherences due to pairs of dipolar coupled protons are correlated with single-quantum coherences resulting in characteristic correlation peaks. Double-quantum coherences between so-called like spins appear as a single correlation peak on the diagonal, while a pair of cross-peaks that are symmetrically arranged on either side of the diagonal reflect couplings among unlike spins. In addition, we exploit the fact that observable double-quantum signal intensities are proportional to D_{ij}^2 or r_{ij}^{-6}, respectively (D_{ij} is the homonuclear dipolar coupling constant, r_{ij} the internuclear distance). Strong signal intensities in the corresponding double-quantum spectrum therefore reveal protons in rather close spatial proximity (*i.e.* distances up to 3.5 Å; Bradley *et al.*, 2009).

The *absence* of a double-quantum auto-correlation peak at 26.4 p.p.m. (13.2 + 13.2 p.p.m.) in the ^{1}H DQ MAS NMR spectrum of OPBA4 (Fig. 9) clearly indicates that the COOH units, unusually, do not form a hydrogen-bonded dimer. Rather, merely 'trivial' double-quantum contacts due to the molecular structure of OPBA4 were found: the cross-peak at 16.0 p.p.m. (9.2 + 6.8 p.p.m.) reflects the spatial proximity of the NH proton and its nearest aromatic proton, while the cross-peak at 12.6 p.p.m. (6.8 + 5.8 p.p.m.) can be attributed to the two different kinds of aromatic protons (one 'regular' aromatic proton and the other one affected by π-stacking). Note the superior spectral resolution of the F2 projection of the ^{1}H DQ MAS NMR spectrum compared with the simple one-dimensional ^{1}H MAS NMR spectrum.

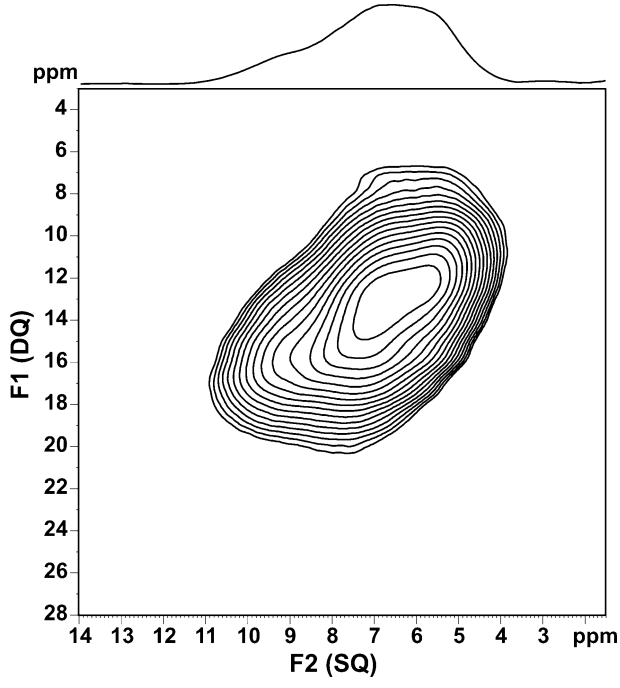


Figure 9
^{1}H/^{1}H correlation measurements of OPBA4.

4. Structures of OPBA2–OPBA3–OPBA4: topological considerations

4.1. OPBA2

The molecules are connected by hydrogen bonds into layers. Each molecule forms four hydrogen bonds, and is connected to three neighboring molecules. Two hydrogen bonds of two carboxylic groups form a centrosymmetric $R_2^2(8)$ synthon linking molecules into dimers. Within the layers the molecules are arranged in a graphite-like honeycomb pattern, while considering the dimers as building units the layer topology corresponds to a square lattice.

4.2. OPBA3

Each molecule forms six hydrogen bonds, two of them building a centrosymmetric $R_2^2(8)$ synthon like OPBA2 molecules in the crystal structure described above. The extended hydrogen-bond network topologically is organized into a three-dimensional double-layered net periodic in two directions. This net consists of two superimposed squared layers interconnected between the four-coordinated vertices, thereby creating five-coordinated nodes. This is a rather specific topological organization, rarely observed for organic molecules; a similar hydrogen-bonded pattern is observed for α-glycine (Dawson *et al.*, 2005). Considering molecular dimers as topological building units, the network can be described as a simple square net, similar to OPBA2.

4.3. OPBA4

Each molecule is connected to four surrounding molecules *via* eight hydrogen bonds forming a layer with the topology of a square lattice. Surprisingly, the carboxylic acid groups do not form a 'default' $R_2^2(8)$ synthon, instead the proton of the carboxylic group forms a hydrogen bond to one of the O atoms of the NO_2 group. The CONH (amide) groups form intermolecular hydrogen bonds, so that a $C(4)$ chain of hydrogen bonds is produced. Each OPBA4 molecule has three amide groups, thus creating three parallel chains linking the molecules into a layer (Fig. 10).

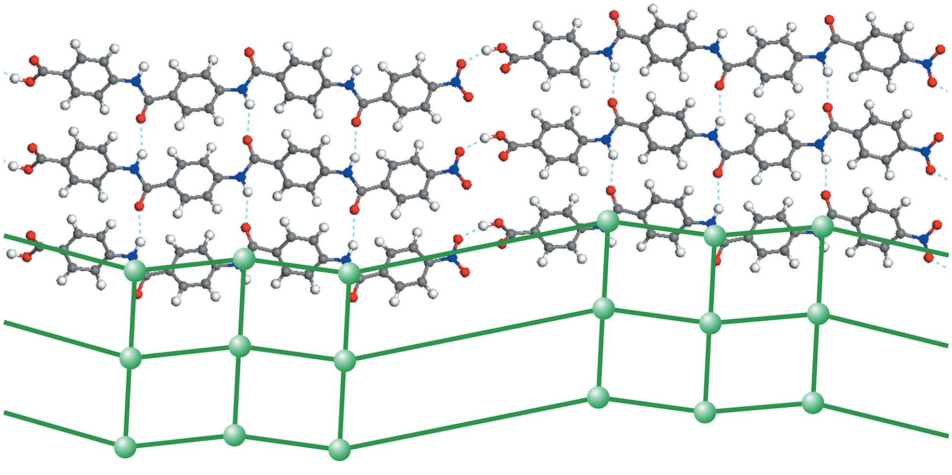


Figure 10
Top view of an OPBA4 layer. The topological lattice is sketched using a green color.

5. Outlook to higher oligomers

The crystallinity of higher oligomers of OPBA (OPBA5, OPBA6, OPBA7) is much lower than that of OPBA4. Electron diffraction patterns recorded using a beam size of 50 nm show broad halos implying the size of the coherent domains in the crystal being far below 50 nm. Crystal structure solution of these materials therefore cannot be carried out using the described ADT procedure. The X-ray powder profiles, on the other hand, are, except for deteriorated crystallinity, very similar to that of OPBA4, suggesting similar molecular packing. Furthermore, solid-state NMR data also show that the packing features must be similar. Therefore, the packing of OPBA4 can likely be used as a model for higher OPBA oligomer packing.

6. Conclusions

Ab-initio structure solution of two related organic compounds was performed from data collected by the automated diffraction tomography (ADT) technique. *Ab-initio* structure solution in *SIR*, although showing a slight distortion in local geometry, produced reasonable structure models both for OPBA3 and OPBA4. Structure refinement against the experimental electron diffraction data (*SHELX*) did not improve the model. Therefore, the final refinement was carried out using energy minimization with the *COMPASS* force field. This novel combination of structure solution and refinement techniques was first tested on the initially known structure of OPBA3 and then applied to the unknown structure of OPBA4. An independent energy-minimization procedure using dispersion-corrected DFT calculations was used for the validation of consistency of all structural models after each step. In the structure determined for OPBA4 an unusual hydrogen-bond scheme from the COOH to the NO_2 group is realised. The absence of the common COOH···HOOC hydrogen-bond synthon is additionally confirmed by solid-state NMR data.

The application of the automated diffraction tomography technique made it possible to solve the structure of a flexible organic molecule '*ab-initio*' despite the high electron-beam sensitivity of the material. Whereas for the smaller homologue OPBA3 the structure could also be solved by X-ray powder diffraction data, this was no longer possible for OPBA4.

The paper demonstrated a beneficial combination of electron diffraction data with force field

structure optimization for molecular crystals. Here, the initial structure model was obtained from ADT data, and the simulation methods were then used for verification and optimization of the model. Simultaneous use of electron diffraction data and packing energy calculations – simulated annealing techniques as they are used for X-ray powder diffraction data – increases the chances of solving the crystal structures of molecular nanocrystals significantly. Moreover, an opposite construction can be arranged – when experimental electron diffraction data is used for validation or selection of structures predicted based on molecular packing energy.

The authors are grateful to Igor Baburin (Dresden, Technical University) for fruitful discussions on the topology of OPBA oligomers and Peter Müller (Massachusetts Institute of Technology) for his generous help with *SHELX* refinement of the structures. The authors would like to thank SFB 625 for financial support.

References

Abbel, R., Frey, H., Schollmeyer, D. & Kilbinger, A. F. M. (2005). *Chem. Eur. J.* **11**, 2170–2176.
Abbel, R., Schleuß, T. W., Frey, H. & Kilbinger, A. F. M. (2005). *Macromol. Chem. Phys.* **206**, 2067–2074.
Antonijevic, S. & Halpern-Manners, N. (2008). *Solid State Nucl. Magn. Reson.* **33**, 82–87.
Bolz, I., Moon, C., Enkelmann, V., Brunklaus, G. & Spange, S. (2008). *J. Org. Chem.* **73**, 4783–4793.
Bradley, J. P., Tripon, C., Filip, C. & Brown, S. P. (2009). *Phys. Chem. Chem. Phys.* **11**, 6941–6952.
Brunklaus, G., Schauff, S., Markova, D., Klapper, M., Mullen, K. & Spiess, H. W. (2009). *J. Phys. Chem B*, **113**, 6674–6681.
Burla, M. C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. & Spagna, R. (2005). *J. Appl. Cryst.* **38**, 381–388.
Chisholm, J. A. & Motherwell, S. (2005). *J. Appl. Cryst.* **38**, 228–231.
Dawson, A., Allan, D. R., Belmonte, S. A., Clark, S. J., David, W. I. F., McGregor, P. A., Parsons, S., Pulham, C. R. & Sawyer, L. (2005). *Cryst. Growth Des.* **5**, 1415–1427.
Gan, Z. (2006). *J. Am. Chem. Soc.* **128**, 6040–6041.
Gan, Z. (2008). *Chem. Commun.* **7**, 868–870.
Gilli, G. & Gilli, P. (2009). *The Nature of the Hydrogen Bond.* IUCr Monographs on Crystallography, p. 23. Oxford University Press.
Gorelik, T., Matveeva, G., Kolb, U., Schleuß, T., Kilbinger, A. F. M., van de Streek, J., Bohle, A. & Brunklaus, G. (2010). *CrystEngComm*, **12**, 1824–1832.
Goward, G. R., Schnell, I. & Spiess, H. W. (2001). *Magn. Reson. Chem.* **39**, S5–S17.
Hirsch, P. B., Howie, A., Nicholson, R. B., Pashley, D. W. & Whelan, M. J. (1977). *Electron Microscopy of Thin Crystals.* New York: Krieger.
Khan, M., Enkelmann, V. & Brunklaus, G. (2010). *J. Am. Chem. Soc.* **132**, 5254–5263.
Kolb, U., Gorelik, T., Kübel, C., Otten, M. T. & Hubert, D. (2007). *Ultramicroscopy*, **107**, 507–513.
Kolb, U., Gorelik, T. E., Mugnailoi, E. & Stewart, A. (2010). *Polym. Rev.* **50**, 385–409.
Kolb, U., Gorelik, T. & Otten, M. T. (2008). *Ultramicroscopy*, **108**, 763–772.
Kolb, U., Mugnaioli, E. & Gorelik, T. E. (2011). *Cryst. Res. Technol.* **46**, 542–554.
König, H. M., Gorelik, T., Kolb, U. & Kilbinger, A. F. (2007). *J. Am. Chem. Soc.* **129**, 704–708.
Kresse, G. & Furthmüller, J. (1996). *Phys. Rev. B*, **54**, 11169–11186.
Kresse, G. & Hafner, J. (1993). *Phys. Rev. B*, **47**, 558–561.
Kresse, G. & Joubert, D. (1999). *Phys. Rev. B*, **59**, 1758–1775.
Krushelnitsky, A., Kurbanov, R., Reichert, D., Hempel, G., Schneider, H. & Fedotov, V. (2002). *Solid State NMR*, **22**, 423–438.
Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989). *J. Magn. Reson.* **85**, 393–399.
Mugnaioli, E., Gorelik, T. & Kolb, U. (2009). *Ultramicroscopy*, **109**, 758–765.
Neumann, M. A. (2011). *GRACE*, Avant-garde Materials Simulation, Freiburg, Germany, see http://www.avmatsim.eu.
Neumann, M. A. & Perrin, M.-A. (2005). *J. Phys. Chem. B*, **109**, 15531–15541.
Olivieri, A. C. (1989). *J. Magn. Reson.* **82**, 342–346.
Orr, R. M. & Duer, M. J. (2006). *Solid State Nucl. Magn. Reson.* **30**, 130–134.
Own, C. S. (2005). PhD thesis, Northwestern University, Evanston, IL, http://www.numis.northwestern.edu/Research/Current/precession.shtml.
Rigamonti, R. (1936). *Gazz. Chim. Ital.* **66**, 174–182.
Saalwächter, K., Graf, R. & Spiess, H. W. (2001). *J. Magn. Reson.* **148**, 398–418.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Streek, J. van de & Neumann, M. A. (2010). *Acta Cryst.* B**66**, 544–558.
Vainshtein, B. K. (1964). *Structure Analysis by Electron Diffraction.* New York: Pergamon Press.
Wang, Y. & Perdew, J. P. (1991). *Phys. Rev. B*, **44**, 13298–13307.
Wilson, A. J. C. (1950). *Acta Cryst.* **3**, 397–398.